Alex Dieker
Kenneth Wu
Luca Perri
Riley T. McArdle

Introduction

For our project, we used Gapminder to find our two quantitative variables of our choosing. Those two variables were "CO_2 emissions per Capita Consumption based" and the "Sustainable Development Index". The 1st dataset contains information that tracks how much CO_2 "consumption-based emissions generated in the production of goods and services according to where they were consumed, rather than where they were produced" (Our World in Data).

The 2nd dataset contains information on the calculation of the development index, which is an efficiency metric made up of 2 other figures known as the "development index" and the "ecological impact index". This index is to help measure different nations' efficiency in delivering human development. Meaning how well can a nation maintain its population's needs while allowing future generations to meet their needs as well.

We hypothesize that there is a negative relationship between CO_2 emissions per capita (consumption-based) and the Sustainable Development Index (SDI). Specifically, as CO_2 emissions increase, we expect the SDI to decrease. The strength of this relationship is expected to be moderate to strong, depending on the country's economic and environmental policies.

The rationale behind this hypothesis is that the Sustainable Development Index penalizes nations for exceeding planetary boundaries—particularly in terms of carbon emissions. Countries that consume more and thus emit more CO_2 per capita are likely to perform worse in sustainability metrics, even if their development index (e.g., life expectancy, education, and income) is high. In other words, high consumption often correlates with high environmental impact, which directly lowers SDI scores.

Cleaning:

https://www.gapminder.org/data/
https://ourworldindata.org/grapher/consumption-co2-per-capita#:~:text=Consumption%2Dbased%20emissions%20attribute%20the,than%20where%20they%20were%20produced.